

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Thastrup, Ole
Tullin, Søren
Boulsen, Lars Kongsbak
Bjørn, Sara Petersen

(ii) TITLE OF INVENTION: Novel Fluorescent Proteins

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Novo Nordisk of North America, Inc.
(B) STREET: 405 Lexington Avenue, Suite 6400
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10174-6401

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/819,612
(B) FILING DATE: 17-MAR-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gregg, Valeta A.
(B) REGISTRATION NUMBER: 35,127
(C) REFERENCE/DOCKET NUMBER: 4594.204-US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 212-867-0123
(B) TELEFAX: 212-878-9655

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGGAATAAGC TTTATGAGTA AAGGAGAAGA ACTTTT

36

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AAGAATTCGG ATCCCTTTAG TGTCAATTGG AAGTCT

36

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 67 nucleotides

006120-07E07500

Pub
C2

60
67

(2) INFORMATION FOR SEQ ID NO:4:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AAGAATTCGG ATCCCTTTAG TGTCAATTGG AAGTCT

36

(2) INFORMATION FOR SEQ ID NO:5:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATTGGTACC AAGGAGGTAA GCTTTATGAG

30

(2) INFORMATION FOR SEQ ID NO:6:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTTTCGTTTT GAATTCGGAT CCCTTTAGTG

30

(2) INFORMATION FOR SEQ ID NO:7:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CATGGCCAAC GCTTGTCACT ACTCTCTCTT ATGGTGTTCA ATGCTTTT

48

(2) INFORMATION FOR SEQ ID NO:8:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTAGAAAAGC ATTGAACACC ATAAGAGAGA GTAGTGACAA GCGTTGGC

48

1. 1990年12月25日，中共中央、国务院作出《关于实行“党政分开”的决定》，指出：“党政分开，不是要削弱党的领导，更不是要否定全心全意依靠工人阶级的方针。而是为了实行和改善党的领导，进一步明确党和政府、党组织和行政组织各自的职权范围，理顺它们之间的关系，使党的领导和政府工作都得到加强。”

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 48 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATGGCCAAC GOTTGTCACT ACTCTCACTT ATGGTGTTCA ATGCTTTT

48

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 48 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTAGAAAAGC ATTGAACACC ATAAGTGAGA GTAGTGACAA GCGTTGGC

48

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGGAATAAGC TTTATGAGTA AAGGAGAAGA ACTTTT

36

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAATCGTAGA TCTTTATTTG TATAGTTCAT CCATG

35

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGAATAAGC TTTATGGATC CCGTCGTTTT ACAACGTCGT

40

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGCGAATTC TTATTATTAT TTTTGACACC AGAC

34

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 764 nucleotides

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTTT ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA ATT CTT 52
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu 15
1 5 10

GTT GAA TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCC GTT AGT GGA GAG 103
Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu 30
20 25 30

GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT ATT TGC ACT 154
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr 45
35 40 45

ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC ACT ACT CTC TCT CAT 205
Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Ser His 65
50 55 60 65

GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT ATG AAA CAG CAT GAC TTT 256
Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe 80
70 75 80

TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG GAA AGA ACT ATA TTT TAC 307
Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Tyr 100
85 90 95 100

AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT GAA GTC AAG TTT GAA GGT GAT 358
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp 115
105 110 115

ACC CTT GTT AAT AGA ATC GAG TTA AAA GGT ATT GAT TTT AAA GAA GAT GGA 409
Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly 130
120 125 130

AAC ATT CTT GGA CAC AAA ATG GAA TAC AAT TAT AAC TCA CAT AAT GTA TAC 460
Asn Ile Leu Gly His Lys Met Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr 150
135 140 145 150

ATC ATG GCA GAC AAA CCA AAG AAT GGC ATC AAA GTT AAC TTC AAA ATT AGA 511
Ile Met Ala Asp Lys Pro Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg 165
155 160 165

CAC AAC ATT AAA GAT GGA AGC GTT CAA TTA GCA GAC CAT TAT CAA CAA AAT 562
His Asn Ile Lys Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn 185
170 175 180 185

ACT CCA ATT GGC GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC 613
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser 200
190 195 200

ACG CAA TCT GCC CTT TCC AAA GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC 664
Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile 215
205 210 215

CTT CTT GAG TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA 715
Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu 235
220 225 230 235

TAC AAA TAA ATGTCCAGAC TTCCAATTGA CACTAAAGGG ATCCGAATTC 764

0064940-071500

Tyr Lys

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 238 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
5 10 15
Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
20 25 30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35 40 45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50 55 60
Ser His Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65 70 75 80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85 90 95
Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100 105 110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115 120 125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
130 135 140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys Asn Gly
145 150 155 160
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
165 170 175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180 185 190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195 200 205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
210 215 220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225 230 235

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 764 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AAGCTTT ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA ATT
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
1 5 10

CTT GTT GAA TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCT GTT AGT 97
 Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser 30
 15 20

GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT 145
 Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe 45
 35 40

ATT TGC ACT ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC ACT 193
 Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr 60
 50 55

ACT CTC TCT TAT GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT ATG 241
 Thr Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met 75
 65 70

AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG 289
 Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln 90
 80 85

GAA AGA ACT ATA TTT TAC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT 337
 Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala 110
 95 100 105

GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA 385
 Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys 125
 115 120

GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA ATG GAA 433
 Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu 140
 130 135

TAC AAT TAT AAC TCA CAT AAT GTA TAC ATC ATG GCA GAC AAA CCA AAG 481
 Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys 155
 145 150

AAT GGC ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT AAA GAT GGA 529
 Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly 170
 160 165

AGC GTT CAA TTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT 577
 Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp 190
 175 180 185

GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACG CAA TCT GCC 625
 Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala 205
 195 200

CTT TCC AAA GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC CTT CTT GAG 673
 Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu 220
 210 215

TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA 721
 Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys 235
 225 230

TAA ATGTCCAGAC TTCCAATTGA CACTAAAGGG ATCCGAATTC 764

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 238 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
 5 10

000120-07E6T560

Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser
 15 20 25 30
 Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
 35 40 45
 Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
 50 55 60
 Thr Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met
 65 70 75
 Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
 80 85 90
 Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala
 95 100 105 110
 Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
 115 120 125
 Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu
 130 135 140
 Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys
 145 150 155
 Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly
 160 165 170
 Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp
 175 180 185 190
 Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala
 195 200 205
 Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu
 210 215 220
 Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 764 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AAGCTTT ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA ATT Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile	49
CTT GTT GAA TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCT GTT AGT Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser	97
GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe	145
ATT TGC ACT ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC ACT Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr	193
ACT CTC ACT TAT GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT ATG Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met	241
AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln	289

000140-07667560

GAA AQA ACT ATA TTT TAC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT 337
 Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala

GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA 385
 Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys

GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA ATG GAA 433
 Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu

TAC AAT TAT AAC TCA CAT AAT GTA TAC ATC ATG GCA GAC AAA CCA AAG 481
 Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys

AAT GGC ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT AAA GAT GGA 529
 Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly

AGC GTT CAA TTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT 577
 Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp

GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACG CAA TCT GCC 625
 Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala

CTT TCC AAA GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC CTT CTT GAG 673
 Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu

TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA 721
 Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys

TAA ATGTCCAGAC TTCCAATTGA CACTAAAGGG ATCCGAATTC 764

- (2) INFORMATION FOR SEQ ID NO:20:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 238 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
 5 10

Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser
 15 20 25 30

Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
 35 40 45

Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
 50 55 60

Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met
 65 70 75

Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
 80 85 90

Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala
 95 100 105 110

Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
 115 120 125

Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu
 130 135 140

Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys
 145 150 155

Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly

000120:07267560

160

165

170

Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp
175 180 185 190

Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala
195 200 205

Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu
210 215 220

Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225 230 235

00549310-071900